ED 155 954 STUDENT EVALUATION OF TEACHING IN MANAGEMENT AND BUSINESS STUDIES. SCEDSIP OCCASIONAL PAPER 2.

1 OF 1 BENINGTON, M. STANDING CONFERENCE ON EDUCATIONAL DEVELOPMENT SERVICES IN POLYTECHNICS, NEWCASTLE UPON TYNE (ENGLAND). 77 34P.

DOCUMENT RESUME

ED 155 954 HE 009 520

AUTHOR	Benington, M.
TITLE	Student Evaluation of Teaching in Management and Business Studies. SCEDSIP Occasional Paper 2.
INSTITUTION	Standing Conference on Educational Development Services in Polytechnics, Newcastle upon Tyne (England).
PUB DATE	77
NOTE	34p.
AVAILABLE FROM	Standing Conference on Educational Development Services in Polytechnics, Ellison Place, Newcastle upon Tyne, England
EDRS PRICE	MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS	*Business Education; Comparative Education; *Effective Teaching; *Evaluation Methods; Foreign Countries; Higher Education; Management; Management by Objectives; *Student Evaluation of Teacher Performance; *Teacher Evaluation
IDENTIFIERS	Great Britain

ABSTRACT

The American system of assessing teacher competence by student rating is evaluated and found to be unsatisfactory for two reasons: a single scale cannot be used to compare teachers doing different tasks; and the scales may not measure 'good teaching' but some other more general social competence. An alternative to the American approach is suggested that contains the management by objectives technique used in industry. The procedure would consist of five stages: (1) the lecturer would work with his subject leader, or deparment head, to identify those changes in knowledge, intellectual and social skills, in attitudes and values they agreed were desireable and appropriate to the course; (2) they would establish criteria for evaluating successful achievement of the objectives, and agree on methods of measurement appropriate to the different objectives; (3) the various forms of assessment would be made; (4) the differences between objectives and achievement would be discussed between the lecturer and his head of department; and (5) the head of a teaching unit from elsewhere in the college would discuss with the lecturer techniques and principles that might help reduce the gap between objectives and performance. (SPG)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED155954

Occassional Paper 2

Student Evaluation of Teaching in Management and Business Studies

M Benington

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Brenda Pollitt

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND TO ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NIE SYSTEM.

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

HE 009520

Price 60p

STANDING CONFERENCE ON EDUCATIONAL DEVELOPMENT SERVICES IN POLYTECHNICS
STANDING CONFERENCE ON EDUCATIONAL DEVELOPMENT SERVICES IN POLYTECHNICS
SCEDSIP **SCEDSIP** SCEDSIP
STANDING CONFERENCE ON EDUCATIONAL DEVELOPMENT SERVICES IN POLYTECHNICS
STANDING CONFERENCE ON EDUCATIONAL DEVELOPMENT SERVICES IN POLYTECHNICS

STUDENT EVALUATION OF TEACHING IN MANAGEMENT AND BUSINESS STUDIES

M. Benington

Bristol Polytechnic

SCEDSIP Occasional Paper 2 (1977)
Editor: Trevor Habeshaw

Published by SCEDSIP,c/o Petras
Newcastle upon Tyne Polytechnic
Ellison Place, Newcastle upon Tyne

Maurice Benington is in charge of the Management Studies Learning Materials Workshop at Bristol Polytechnic, which is supported by the Foundation for Management Education.

ABSTRACT

The American system of assessing teacher competence by student rating is unsatisfactory for two reasons: a single scale cannot be used to compare teachers doing different tasks; and, the scales may not measure 'good teaching' but some other more general social competence.

A satisfactory alternative to the American approach would include provision for measuring the lecturer's success at achieving the objectives he has chosen.

STUDENT EVALUATION OF TEACHING IN MANAGEMENT AND BUSINESS STUDIES

As Polytechnic and Technical Colleges get bigger and become more bureaucratic, pressures for more objective forms of staff appraisal are likely to increase, and because of the presumed relevance of their experience of performance appraisal in industry, Management and Business Studies teachers may be asked to co-operate in the setting up of staff appraisal schemes. Since an important part of a lecturer's job is teaching, staff appraisal schemers may wish to measure lecturers' teaching ability, (as was suggested by the Prices and Incomes Board in 1968). This article is intended to indicate some of the problems associated with student evaluation of teaching and to suggest some features which appraisal schemes should contain.

Many colleges in the United States make use of teacher evaluation; Bryan (1968) for

example, found that 49% of large colleges offering Bachelor's Degrees did so, and it is largely from the American experience that the substance of this article is drawn. The most common method of measuring teacher effectiveness, in the United States, is by student evaluation, but assessment by Chairmen of Departments has also been tried.

One of the defects of rating teacher effectiveness by Departmental Chairmen is that of the subjectivity of one individual's rating; the rating may be affected by personal bias toward leniency or severity; by Halo Effect - in which the presence of one favourable characteristic affects the rater's judgement of other characteristics; and by personality traits, such as authoritarianism. (A summary of these and other factors is given by Anderson and Hunka (1963). Assessments made by a group of students, on the other hand, are unlikely to be all systematically biassed in the same direction, and so do not suffer from this defect.

A second defect of Chairman's rating is that his perception may be different from that of students.

As Sorenson and Gross (1967) wrote:

> *"teachers are bound to be regarded differently by persons with varying concepts of the teacher's role".*

But the extent of the agreement between students and Chairmen is an empirical question which is not fully settled. Blackburn (1975) found that a professor's peers agreed with his students on his teaching effectiveness. Gromisch (1972) found no correlation between students and Chairmen when professors were rated on a 38-item questionnaire. Sagan (1974), found no correlation between students and staff, nor between students and chairmen; but there was agreement between staff and chairmen - a finding consistent with students having different perceptions of a good teacher from those of college staff.

If Chairmen and students do not reliably agree on an assessment, the question is raised, whose opinion should be sought?

A frequent reply is, that since students are consumers of instruction, they should be the ones to judge it. This fancy, together with the problem of bias in a single rater, have meant that most teacher evaluations are based upon student assessment.

If student evaluations were to be used to help a lecturer to improve his teaching skill, it might be more helpful to the lecturer if students made free comments on what they thought was wrong with his delivery, and what they thought he could do to put it right. But where evaluation is to be made the base of appraisal, it would seem at first glance to be more appropriate to compare teachers with one another on a standard scale* Whether this is a legitimate activity will be discussed below.

Beneath the plan to compare teachers on a standard scale lies the implicit assumption that lecturers are doing something comparable, and that each item of the questionnaire applies with equal force to each lecturer; that for all classes, all syllabuses, all objectives, all

* A fairly typical scale is reproduced in the Appendix.

student abilities, all types of students, teachers shall equally enthuse, participate, explain clearly and be sensitive to their students. This assumption of comparability between lecturers is false; different lecturers have different objectives, teach different students, use different materials; and because of these differences, require different teaching techniques. Evidence for this comes from a number of studies.

Ramsden (1975), found that Arts and Social Science students thought that sensitivity to students was an important teacher characteristic; but Science and Engineering College students valued this sensitivity significantly less.

He also found that Science students valued lecturers who were clear and understandable significantly more than Arts students. And it was of greater importance to Science students than to Arts students that the lecturer proceeded at a

suitable pace for the class to take notes.

In an unpublished study by the author, an attitude survey was conducted among Day students and Evening students working for the same professional exam. The lecturer had adopted an aggressive, bullying, 'Theory X' approach. Significantly more of the evening group agreed with the statements:

"The style of delivery is suitable for working adults"

"The lecturer treats us like responsible adults"

An explanation for the difference in response between the two groups was offered by the evening students; as evening students were much harder pressed by other demands of their time than were the day group, they found that the lecturer's urging and driving was helpful, and they resented it much less.

Differences in teaching behaviour are called for not only by differences in the nature of the

course and the circumstances of the students, but also by differences in teaching objectives. Management and Business Studies courses may pursue a wide variety of objectives; a few possibilities are listed in Table 1.

An important feature of these different objectives is that they do not require similar types of learning. And as the Belbins (1969) have pointed out different types of learning require different teaching techniques. Some teaching techniques which are more appropriate to different types of learning, are tabulated. See Table 2.

Table 1 Possible Teaching Objectives for a Management or Business Studies Course

Type of Learning	Possible Objectives
M	To present an accurate and complete body of knowledge for use during the students career.
A	To develop in students a non-dogmatic, tolerant balanced approach to the consideration of controversial issues.
S	To develop the skills of evaluating and learning information as they are required.
A	To develop a capitalistic, entrepreneurial business-oriented ethic.
C, A	To develop in students the understanding of business organisations as social systems, to enable them to develop equitable solutions to management problems.

KEY: M MEMORY A ATTITUDE S SKILLS C COMPREHENSION

Table 2:

Teaching Technique	Type of Learning	
	C	M
Ensure plenty of practice	-	++
Revise periodically	-	++
Prevent all error in early stages	-	++
Avoid frequent changes of task	-	++
Feed information slowly	++	++
Learn by exercise rather than by words	+	-
Concentrate on the general rather than on the particular	++	-
Break down the task into parts	-	++
Focus teaching on key points	++	-
Teach individually, if possible	+	-

(Adapted from Belbin, 1969)

These tables indicate, for example, that if teaching is directed to getting students to amass a certain body of knowledge, then the learning activities that lecturers should encourage are completely different from those that are important to understanding a concept or theory. It would seem to follow from this argument, that one cannot appraise teachers on standard forms, unless they are teaching comparable students, in comparable situations, and are attempting to teach the same objectives at the time of evaluation.

It has been argued (Bantock, 1961) on logical grounds alone, that there can be no universally valid criteria of teaching - a conclusion entirely consistent with the conclusion reached above. Bantock's argument has been criticised by Martin (1963), who argued that it was fallacious; but where the philosophers may disagree, there does seem to be a substantial body of opinion deriving from empirical research that would support Bantock's conclusion. In studies based on the replies of 7,000 students (Sheffield, 1974) identified and studied

24 "best" lecturers; he concluded that not all good teachers are good for all students; and that students from different faculties look for different things. Sharples and Woodman (1972, note that the attributes of successful teachers differ as between different teacher tasks. McKeachie (quoting Bendig, 1955) writes:

> " *in measuring teacher competence, all teaching goals should be assessed. A teacher who is effective in achieving one goal is not necessarily effective in achieving others.*"

One implication of these conclusions might be to abandon the attempt to create a standard assessment scale by which to compare all lecturers, and to replace it by individually constructed scales, intended to measure how well a lecturer is achieving his chosen objectives. A very interesting account of such a method is reported by Boud and Turner (1974). In this study they drew up a list

of objectives for a course taught by one author. They then asked the students, how far were these objectives desirable in an ideal course; how far were they seen to be objectives of the present course; were the objectives being achieved, and did they think that the objectives would be examined. The study demonstrated that there were some wide differences between students' perceptions of the objectives of an ideal course and the extent to which these objectives were being achieved on this course. The study was found to be a useful method for uncoverning problem areas in that course.

To abandon standard evaluation forms as Boud and Turner have done, does mean that simple mechanical appraisal procedures are abandoned. But the simple procedure may be replaced by a far more helpful one that has resemblances to the procedures known as Management by Objectives, used in industry for supervising ones subordinates. If we adapt the Boud and Turner procedure by a small extension in the direction of Management by Objectives, a procedure of

staff evaluation can be outlined that is consistent with a systematic approach to teaching and with accepted principles of good management, a procedure that is likely to engage the support of teaching staff, and one that makes use of student opinion, where student opinion is probably a valid measure of lecturer effectiveness.

The procedure would consist of five stages.

In the first stage, the lecturer would work with his subject leader, or Department Head, to identify those changes in knowledge, intellectual and social skills, in attitudes and values they agreed were desirable and appropriate to the course, and express them as behavioural objectives.

Secondly, they would establish criteria for evaluating successful achievement of the Objectives, and agree on methods of measurement appropriate to the different objectives.

On such objectives as, "To create a favourable attitude among students to the subject", and "To help students to see the relevance of the subject to their future vocations", student opinion would seem to be an appropriate measure. On the other hand, for such an objective as, "To enable students to solve simple simultaneous equations in two variables", objective examinations would seem to be more suitable.

In the third stage, the various forms of assessment would be made.

Fourthly, the differences between objectives and achievements would be discussed between the lecturer and his Head of Department.

Finally, the Head of perhaps a Teaching Unit from elsewhere in the College, would discuss with the lecturer lecturing techniques and principles that might help to reduce the gap between objectives and performance.

A second cause for disquiet in the use of ordinary evaluation scales, is that we can't be sure that

they do measure good teaching. The items for a scale usually consist of statements of desirable teacher behaviour; and in use, students assess the extent to which the lecturer displays the behaviour. The scale constructor arrives at statements of 'desirable' behaviour either by asking students what they think is desirable; or by deriving statements from Learning Theory; or simply by representing his own value judgements, such as - "good teachers are democratic, involve students, and dress well".

The trouble with each of these approaches is that the scale constructor has still to show that they are valid. That is to say, that teachers who carry out these behaviours, are in fact good teachers. This brings us to the nub of the problem: What is a good teacher? and Do these scales measure good teaching behaviours?

There is no way by which the validity of these questionnaires can be demonstrated until good teaching has been defined.

It is not enough to assert that: "students recognise good teaching", because that statement itself requires validation, which is not possible without the required criterion, i.e., a statement defining "good teaching". Moreover, there are good reasons for thinking that students do not always recognise good teaching when they meet it, and that they sometimes judge appalling teaching to be good. Solomon (1963), found that a teacher who promoted the deepest interest in his subject, and whose students learnt most, and whom experts thought most highly of, received an unenthusiastic response from his students. Natuflin et al (1973), found that a charismatic presentation by an actor that was intentionally irrelevant, conflicting and meaningless, received quite favourable responses from his students on such items as "Did he put his material across in an interesting way"? and "Did he present his material in a well-organised form?"

There can be little doubt however that students' evaluations do seem to reliably measure something. As Remmers (1963), concludes, student

evaluations based on more than 25 students are as reliable as the best psychological tests. In his thorough review of student evaluation, Flood Page (1974), shows that these ratings are not substantially affected by irrelevancies, such as students' grades or mere showmanship by the lecturer.

Reliable they may be: but whether student evaluations reliably measure 'good' teaching is another matter. A consideration of some items on a typical scale reminds one forcibly of some approaches to the study of leadership. According to Adair (1967), the successful leader takes actions to meet the needs of the individuals in the group, to meet group needs, and to meet the needs of the task.

Some items from questionnaires falling into these categories are shown overleaf.

Table 3 Questionnaire Item

Task Needs	Is clear and understandable in his explations
	Makes good use of examples and illustrations
Group Needs	Inspires class confidence in this knowledge of the subject
	Is enthusiastic
Individual Needs	Pays attention to students
	Shows an interest in individuals

What these items appear to be measuring then, is not something identifying characteristics of teachers, but some characteristics common to many social situations. The same point is made by Gibb (1955), who having noticed the similarity between teacher behaviour and some other types of behaviour in individual-group interaction based his teacher evaluation form on a Leadership Behaviour Description Questionnaire. Considerable similarities were demonstrated between college teachers and other leaders.

Perhaps such scales are measuring Leadership, Social Competence or Charisma; or alternatively, they could be measuring teaching; but it remains for the users to demonstrate that this is what they do measure. To argue that the scales measure what teachers do, and are, therefore measuring teacher behaviour, is to overlook the fact that teachers perform many roles, and it may not be the pedagogical role that is being assessed by the scale.

It would be possible to throw up many suggestions for criteria of good teaching but there would be no means of distinguishing between good and bad suggestions and no means of getting agreement between educators as to what constituted good teaching until agreement has been reached on the nature of the concept. That is to say, we cannot expect agreement until we have a theory of teaching by which we can judge our choice of criteria. This is not an empirical matter; it is something for educators and philosophers to arrive at through the exercise of their values and judgement.

The validity of student assessment has been approached from a more limited perspective that the one outlined above. Taking the view that examination results are a criterion of good teaching, Marsh et al. (1975), compared the average performance of students pursuing the same course, assessed by the same final examination, but taught by different instructors. The students' final grades were compared to a number of teacher variables, By this means,

they established the validity of some items taken from a student evaluation scale, against the criterion of examination success. Valuable though it is, the method is weak in that we cannot use it for measuring most teaching goals. The examination was used to measure the acquisition of factual knowledge, but it appears that other important educational objectives, such as the development of intellectual skills, or the acquisition of academic values were not assessed.

If examinations are to be used as a criterion of teaching success, it is surely necessary to demonstrate that the examinations measure all significant educationals goals.

The consequences for a department of using student-based teacher-evaluation as a basis for staff appraisal will depend upon many organisational factors. These would include such matters as the present relationship between staff and students, between staff within the department, the lecturer's tenure,

the prospects for promotion, and the diffusion of power throughout the department. It seems to the author that it would be unwise to introduce a formal system of appraisal of teaching into British Colleges, based on the typical American method of standard assessment by students, with all its inadequacies, without first considering the probable effects upon the social system of the users' departments. The nature of these effects, would be the subject for another article.

REFERENCES

ADAIR, J E — 1967 "New Trends in Leadership & Management Training". Royal United Service Inst. Journal 112 (p. 315-23)

ANDERSON C C & HUNKA S M — 1963 Teacher Evaluation: some problems and a proposal. Harvard Educational Review 33 74-95

BANTOCK G H — 1961 Educational Research : a criticism Harvard Educational Review 31 264-280

BELBIN M R — 1969 CRAMP: an Introduction to CRAMP. Training International, 4

BLACKBURN R T — 1975 An assessment of faculty performance: some correlations between administrator, student and self ratings. Sociology of Education 48 (2) 242-256

BOUD D J & TURNER — 1972 A Method for the Evaluation of a Traditional Lecture Course. IN Research into tertiary science education: a selection of papers from the Conference on Tertiary Science Education Dec 1972 Soc. for Research into H.E. 1974 p 67-74

BRYAN R C — 1968 Student Rating of Teachers- improving College and University Teaching 16 200-202

FLOOD-PAGE C 1974 Student Evaluation of Teaching: the American experience.
London Society for Research in Higher Education E2:60

GIBB C A 1955 Classroom Behaviour of the College Teacher.
Educational and Psychological Measurement 15,
254-263

GROMISCH D S et al 1972 A comparison of student and departmental chairmen's evaluations of teaching performance.
Journal of Medical Education, 47,
281-284

JOHNSON H C et al 1975 Higher Education 4, 1975
173-199 Part 1, and
Higher Education 4, 1975
273-303 Part II

MARSH H W FLEINER H & THOMAS C S 1975 Validity and usefulness of student evaluations of instructional Quality.
Journal of Educational Psychology
1975, 67 (6) 833-839

MARTIN J R 1963 Can there be universally applicable criteria for good teaching.
Harvard Educational Review 33
484-491

NATUFLIN D H WARE J E Jr DONNELLY F A 1973 "The Doctor Fox Lecture": A paradigm of Educational Seduction".
Journal of Medical Education, 48,
630-635

REFERENCES

RAMSDEN P — 1975 Polytechnic Students' Expectations of their teachers, and the use of a student feedback quesionnaire. Higher Education Bulletin, Vol 3 No 2, Spring 1975

REMMERS H H — 1963 (in Gage, N L (1963) Chapter 7 - "Rating Methods in Research on Teaching" Handbook of Research on Teaching, Rand McNally & Co. Chicago

SAGAN H B — 1974 Student faculty and department chairmen ratings of instructors: who agrees with whom? Research in Higher Education 2 (3) 265-272

SHARPLES D & WOODMAN P — 1972 Multiple and Task Specific Models in the Assessment of Teaching. Durham Research Review, Vol vi, No 28 1972 Page 627

SHEFFIELD E F — 1974 Teaching in the Universities: no one way. Montreal & London - McGill - Queen's University Press, xiv, and p. 252

SOLOMON D BEZDEZ W E & ROSENBERG L — 1963 Teaching Styles & Learning. Brooklin Mass, Centre for the Study of Liberal Education for Adults.

LECTURER ASSESSMENT SCALE

This scale is to be used in a research project concerned with the improvement of teaching methods. I would be grateful if you would answer each question with reference to this lecture. We are not interested in the lecturer's presentation on any occasion other than this one.

Please indicate the extent of your agreement with each statement by placing a number in each box.

Example:

Suppose that one of the items was

18 Spoke in a clear voice [ ]

If you thought the lecturer had a very clear voice you would strongly agree with the statement: in this case put 5 in the box. Like this: [5]

On the other hand if you thought that the lecturer mumbled so that you could scarecely hear him, you would strongly disagree with the statement. In this case you would put 1 in the box. [1]

If you can't make up your mind what you think, or don't know, put 3 in the box.
Different measures of agreement can be expressed by using 4 or 5.
Different measures of disagreement can be expressed by using 2 or 1.

If you

Strongly agree	put	5
Agree	put	4
Undecided	put	3
Disagree	put	2
Strongly disagree	put	1

If you think the item was irrelevant or if you don't know put 0

	THE LECTURER	PLEASE LEAVE BLANK COLUMN
1	Presented his subject so that I could understand it	☐
2	Made his subject interesting	☐
3	Presented his material clearly and logically	☐
4	Thoroughly prepared the lesson	☐
5	Indicated what we should read to "fill in" the lecture	☐
6	Made it plain how this lecture fitted into his lecture series	☐
7	Kept to the point, without wasteful digression	☐
8	Provided a framework for further study	☐
9	Organised his blackboard work clearly	☐
10	Used visual aids where appropriate	☐
11	Used effective, clear visual aids	☐
12	Tried to ensure that the students understood him	☐
13	Used teaching techniques other than continuous speech	☐
14	Made it clear what students had to learn	☐
15	Helped us to assess whether we understood what was being taught	☐

PLEASE LEAVE BLANK COLUMN

16 Answered questions helpfully ☐

17 Gave references for further or remedial study ☐

18 Spoke clearly and could be heard comfortably ☐

You may wish to make some comments of your own in the space below:

LECTURER'S NAME .....

COURSE (e.g. HND, BA Business Studies, etc)

.....

.....

SUBJECT .....